第41卷 第10期
2024年10月

干 旱 区 研 究
ARID ZONE RESEARCH

Vol. 41 No. 10
Oct. 2024

基于决策树的天山冰湖提取方法研究

李梦帆[1]， 郑江华[1,2]， 钱安良[1]， 李家辉[1]， 阿迪力江·帕尔合提[1]，
王 哲[1,2]， 马丽莎[1,2]， 王 南[1,2]

（1. 新疆大学地理与遥感科学学院，新疆 乌鲁木齐 830017；
2. 新疆绿洲生态自治区重点实验室，新疆 乌鲁木齐 830017）

摘 要：天山位于亚欧大陆中部，是现代冰川的主要分布区之一，该地区冰川融水形成了数量多且分布广泛的冰湖。冰湖是气候变化的重要指示器，也是中国西北干旱与半干旱地区重要的地表水及地下水供给来源。由于地形因素和地物光谱特征的影响，使用单一的水体指数进行遥感影像的冰湖提取时，难以较好地区分出冰湖、山体阴影和积雪。本研究以天山地区为研究区，基于Google Earth Engine云平台，以Landsat 8遥感影像为数据源，根据冰湖的空间位置（缓冲区范围）、地形特征（坡度、高程）以及光谱特征，构建了冰湖决策树提取方法，并与NDWI（归一化水体指数）、MNDWI（改进的归一化水体指数）阈值法进行了精度比较。结果表明：决策树法能够有效减小山体阴影和积雪影响，更有效地提取冰湖信息，提取结果总体精度为89.14%，Kappa系数为0.783，F1分数为87.85%。结合了空间位置、地形特征和光谱特征的决策树方法为冰湖的动态监测与研究分析提供了一种较为高效的提取方法。

关键词：冰湖提取；决策树；Google Earth Engine；天山

冰湖是由于高山冰川融水在地表洼地汇集而形成的天然湖泊，冰湖的扩展和枯萎与冰川、气候变化有密切关系[1]。冰湖数量和面积的增减是区域气候变化的重要指示剂，如何快速解译冰湖信息、建立冰湖的动态监测系统是气候变化背景下的重要问题。亚洲高山区作为第三大陆地冰冻圈，分布着大量的冰川积雪，是世界上发育山地冰川最多的中纬度地区[2]，其冰川融水是周围干旱半干旱区的重要补给来源[3]。其中天山山系深居亚欧大陆内部，是中国现代冰川的主要分布区之一，其丰富的冰川资源对于新疆社会经济发展、自然资源利用及生态文明建设等提供了有力的水资源保障[4]。

实际研究中，冰湖提取过程可分为两个步骤，首先根据冰湖的地形特征确定其大致分布范围，即冰湖范围的粗筛选；然后利用水体指数对冰湖进行精提取。在确定其分布范围时，则需要从冰湖的补给来源、产生机理等方面对冰湖定义做出解释。在最初的研究中，有学者将冰湖定义为受到冰川作用而形成的湖泊[5]，但是缺乏了对于作用时间和湖水来源的具体描述。随着对冰湖编目和冰湖演化与冰川变化耦合研究工作的进一步展开，冰湖被定义为“末次冰期冰盛期以来在受冰川侵蚀作用形成的洼地中积水而形成的天然水体”[6]，陈晨等[7]综合了现代过程和融水对冰湖形成作用角度提出冰湖是指末次冰期以来冰川在进退过程中形成湖盆，并由现代冰川融水或降水补给的湖泊；姚晓军等[6]对冰湖的定义“冰湖是以现代冰川融水为主要补给源或在冰碛垄洼地内积水形成的天然水体”更加强调冰川在冰湖形成和变化中所起到的作用。

水体的总体反射率大约在3%左右，主要集中在380~760 nm的可见光波段，在750 nm处，水体的反射率显著衰减[8]。目前，普遍应用的水体指数有NDWI（归一化水体指数）、MNDWI（改进的归一化水体指数）等，以NDWI的阈值提取方法，具有简单快速的特点。Wang等[9]利用了Landsat卫星系列数据，计算高亚洲地区NDWI图像，经过人工目视解译

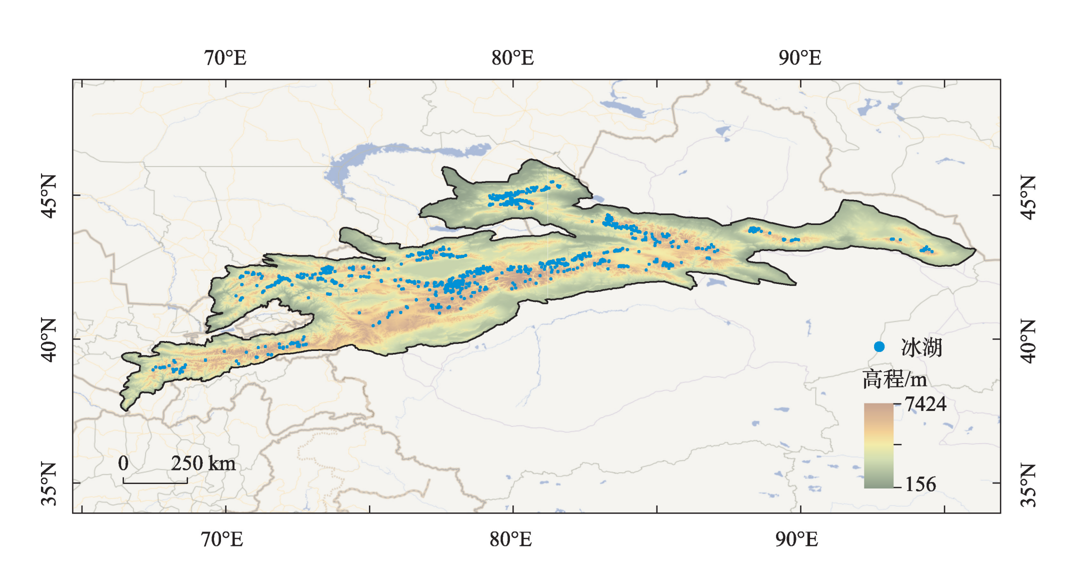
收稿日期：2024-03-26； **修订日期：**2024-06-25
基金项目：国家级大学生创新研究项目（202210755003）；第三次新疆综合科学考察项目（2021xjkk1001）资助
作者简介：李梦帆（2002-），男，主要研究方向为资源环境遥感. E-mail: l_mengfan@126.com
通讯作者：郑江华. E-mail: zheng.jianghua@xju.edu.cn

对冰湖边界矢量化后得到1990年和2018年高亚洲冰湖编目。当冰湖研究区范围过大时，考虑到全局阈值方法会造成的不同区域的冰湖提取精度不一，骆剑承等[10]提出了"全局-局部迭代分割法"，包括NDWI计算、全局分割、局部迭代分割及专题图生成4个步骤。随着对冰湖提取研究的不断深入，冰湖提取也不仅限于单一阈值方法。Govindha等[11]通过手动设置NDWI阈值，并叠加DEM数据进行喜马拉雅山锡金地区冰湖的识别。Mitkari等[12]基于Sentinel-2影像，通过阈值法与Sobel边缘检测法结合，实现了对甘戈特里冰湖的半自动提取。闫斌等[13]基于NDWI-NDSI组合法使得冰湖的错分现象得到改善。此外，还有学者对常见的提取算法如人工阈值法、OTSU迭代法、U-NET、双峰迭代法、C-V迭代法、决策树法进行了对比，验证了OTSU迭代法、C-V迭代法在喜马拉雅地区的可行性[14-15]。冰湖多分布于高原山地地区，山体阴影具有与水体相似的光谱曲线[16]，且受到周围冰川积雪影响，冰湖的提取工作较为复杂。Li等[17]在"全局-局部迭代分割法"的基础上，引入了DEM数据生成山体坡度和阴影信息，一定程度上消除了喜马拉雅山地区山体阴影的影响。该方法在一定程度上实现了冰湖的自动提取，但对于与背景信息具有相似光谱特征的微小冰湖，则会造成特征混淆[18]，影响自动识别精度。近些年来，机器学习在地物解译中得到广泛应用，Jain等[19]使用SVM(支持向量机)法对喜马拉雅山不丹地区冰湖进行了半自动化检测。Veh等[20]基于Landsat数据使用RF(随机森林)算法对1988—2016年喜马拉雅山冰湖进行识别，总体精度可达91%。Bhardwaj等[21]同样根据Landsat影像，结合湿度指数和模糊聚类算法提取了Shaune Garang冰川附近的冰湖，提取精度可达99%。但基于机器学习的冰湖提取过于依靠训练样本的可靠性，且耗费大量算力，不适用于大范围地区冰湖的提取。

本文基于GEE平台获取2022年天山地区Landsat 8遥感影像数据，利用湿度指数提取冰川，在空间位置和地形特征方面使用坡度、高程、缓冲区分析等确定冰湖分布范围，随后根据冰湖的光谱特征提出NDWI-SWI组合阈值法进行冰湖信息提取，最后经过检查修正和面积筛选得到天山地区冰湖数据。本文基于遥感大数据云计算平台，建立了以空间位置、地形特征与光谱特征结合的决策树水体提取方法，为冰湖半自动化高效提取的实现提供了一种新的方法选择。

1 研究区概况

天山地区(图1)位于亚洲中部(37°~47°N，67°~97°E)，西起乌兹别克斯坦的克孜尔库姆沙漠以东，跨越哈萨克斯坦、吉尔吉斯斯坦以及中国新疆，东至哈密市以东的戈壁中，主要分布在中国、哈萨克斯坦、吉尔吉斯斯坦、乌兹别克斯坦境内，东西最长达2500 km，南北宽250~350 km，山地面积约60×



注：底图采用自然资源部标准地图制作，审图号为GS(2021)5453号，对底图边界无修改。下同。

图1 研究区概况图

Fig. 1 Overview of the study area

10^4 km²。天山地区属于温带大陆性干旱气候，主要表现为夏季炎热，冬季寒冷，年温差较大。并且具有明显的垂直地带性分布特征。天山地区地处亚洲中部地区，远离海洋，空气干燥，降水不足，降水量随海拔升高呈现递增趋势，且地形复杂，形成了分布不均匀的降水格局[22]。在全球变暖的气候背景下，冰湖变化一方面可以影响水资源补给，另一方面也存在泥石流等自然灾害的潜在风险[23-27]。

2 数据来源与方法

2.1 平台及数据源

GEE（Google Earth Engine）是由谷歌、卡内基梅隆大学和美国地质调查局联合开发的遥感云计算平台。该平台具有丰富的遥感数据源、封装API函数以及快速处理能力，极大地提高了大尺度研究区的数据处理及分析效率。本研究使用的影像数据及地面高程数据均在该平台导入并处理。

本文以GEE平台Landsat 8地表反射率数据集（LANDSAT/LC08/C02/T1_L2）为遥感数据来源，空间分辨率为30 m。为减少云量及积雪等因素对冰湖提取的影响[28]，选取2022年9—10月影像，对63景影像进行逐景筛选，使得每景影像的云量均为研究时段对应行列号最低，平均云量为2.1%，最后进行镶嵌裁剪，得到研究区影像。

地面高程模型使用NASADEM（NASA/NASADEM_HGT/001），NASADEM是STRM数据的再处理，空间分辨率为30 m（https://earthdata.nasa.gov/），本研究基于该数据提取研究区的高程、坡度等地形因子，进而对冰湖的空间范围进行筛选。

2.2 研究方法

2.2.1 *决策树法* 决策树法的基本思想是通过构建一定规则从原始影像中进行逐层的掩膜处理，以实现对目标的分离。本研究通过高程、坡度、缓冲区范围、水体指数阈值等规则进行冰湖的提取，主要分为3个步骤。首先通过空间位置（缓冲区范围）、地形特征（高程、坡度）确定冰湖的分布范围；在此基础上利用NDWI-SWI两种水体指数组合进行冰湖提取，最后在ArcGIS中进行河流剔除、冰湖面积筛选并检查修正，生成研究区冰湖数据。具体技术流程如图2所示。

2.2.2 *冰川提取* 缨帽变换是Kauth与Thomas于1976年根据地物的多维光谱分布特征提出的图像

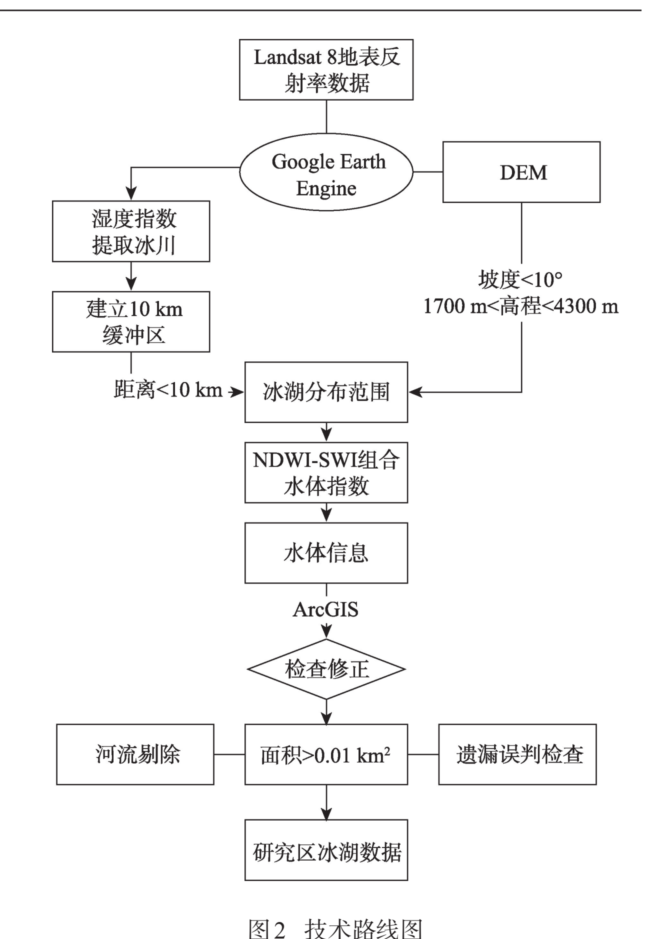


图2 技术路线图

Fig. 2 The technical flow chart

变换经验线性方法[29-30]。经过缨帽变换后产生的亮度、绿度、湿度等分量与地物特征联系更加密切，能够突出特定信息。其中第三分量湿度分量反映了地物的水分信息。周远刚等[31]利用Wetness指数对冰川进行了识别。本研究在实验中发现当湿度指数>0时可以将天山地区的冰川有效提取出来，为后续建立冰川缓冲区，进而为确定研究区冰湖的空间分布范围提供数据来源。

2.2.3 *冰湖空间位置与地形特征* 由于冰湖是以冰川融水为主要补给来源在洼地积累形成，因此，冰湖具有一定的分布范围和地形特征。吴坤鹏[22]在天山冰湖变化及其影响的研究中提出天山地区冰川湖泊分布在海拔高度为1700~4300 m的范围内；王欣等[32]在喜马拉雅山地区冰湖边界提取中，将坡度阈值设置为5°能较好地区分冰湖和山体阴影；杨成德等[33]综合国内外学者研究，将0.0036 km²作为中国西部冰湖编目的最小面积，将冰川10 km范围内作为冰湖分布区。本研究综合已有研究方法及天山地区地形特点，将冰湖空间位置与地形特征定义为H：

$$H=(D\leqslant 10\ \mathrm{km})\wedge(S<10°)\wedge(1700\ \mathrm{m}\leqslant E\leqslant 4300\ \mathrm{m})\wedge(A>0.01\ \mathrm{km}^2) \tag{1}$$

式中：D表示冰湖到冰川的距离(km)；S为坡度(°)；E为海拔(m)；A为冰湖面积(km²)。通过提取到的冰川数据建立缓冲区，再利用DEM数据对坡度、海拔进行计算和阈值选取以建立掩膜，通过掩膜叠加确定冰湖的分布范围，在数据检查修正环节筛选符合面积要求的冰湖。

2.2.4 NDWI（归一化水体指数） McFeeters[34]利用近红外波段和绿波段提出了归一化水体指数(NDWI)，定义如下：

$$\mathrm{NDWI}=\frac{Green-Nir}{Green+Nir} \tag{2}$$

式中：*Green*和*Nir*分别为绿波段和近红外波段的反射率。

2.2.5 SWI（阴影水体指数） 陈文倩等[35]发现水体在蓝绿波段下降较慢，几乎呈平行状，而山体阴影在蓝绿波段则下降较快，并构建阴影水体指数SWI，应用于西北山区，可以较好地区分水体和阴影。SWI定义为：

$$\mathrm{SWI}=Blue+Green-Nir \tag{3}$$

式中：*Blue*、*Green*、*Nir*分别为蓝、绿、近红外波段反射率。在光谱特征提取方面，一些学者在进行水体提取时发现NDWI指数在提取水体信息的时候会将山体阴影误提[36-38]。对于复杂山区而言，单一的水体指数方法提取效果并不理想。因此，加以SWI指数进行光谱特征方面的第二次筛选，将山体阴影部分去除。根据研究区实际情况，在经过多次实验后发现，当水体指数阈值组合为$0.12<\mathrm{NDWI}<1.3$及$\mathrm{SWI}>0.02$时，能大幅去除山体阴影影响并提取水体。

3 结果与分析

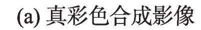
3.1 提取结果与分析

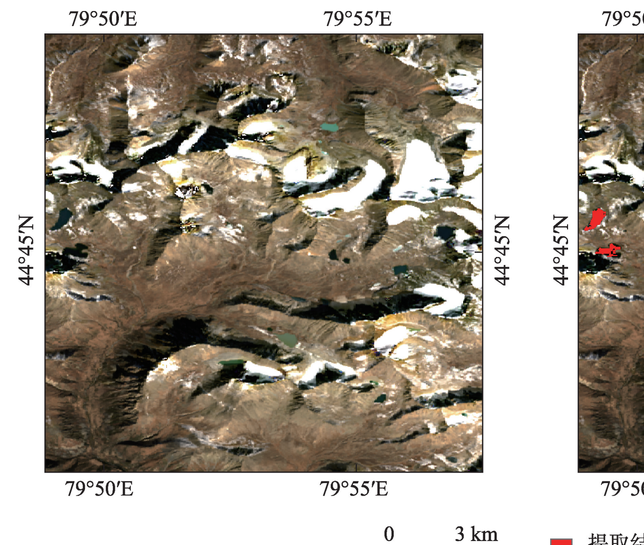
因研究区范围较大，为具体展示提取效果，选择冰湖分布较多的局部区域，经纬度范围及提取结果如图3所示。为进行数据提取效果比较，本文使用决策树法、NDWI法、MNDWI法[39]对研究区影像进行处理，提取结果对比图如图4所示。图4a为真彩色合成影像，其中NDWI、MNDWI阈值法由于整个研究区范围过大，灰度直方图不具有明显的波峰波谷特征，因此，均选取局部区域范围后再进行阈值分割处理。

在所示区域内，NDWI阈值法虽然能够提取大部分冰湖(图4c)，但同时也提取了大量积雪和少量山体阴影。MNDWI阈值法受积雪和山体阴影的影响极大，且对于冰湖的提取效果不佳(图4d)。决策树法则能够准确地将冰湖信息提取出来，并完全排除了积雪与山体阴影的干扰，提取效果最佳(图4b)。

3.2 提取精度验证

为检验决策树法提取冰湖的准确性，本文基于研究区原始影像、Google Earth地图，在天山地区随





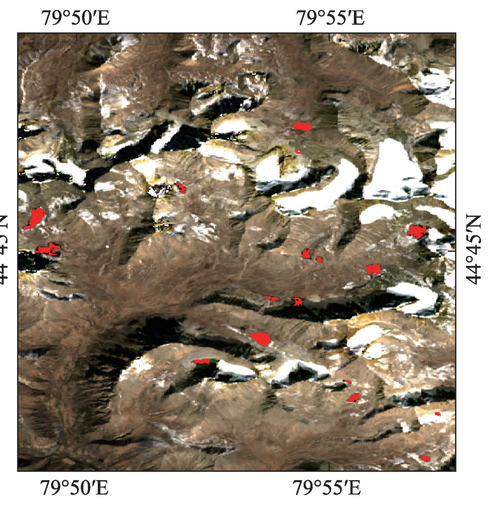


图3 局部地区冰湖提取结果

Fig. 3 Results of glacial lake extraction in local area

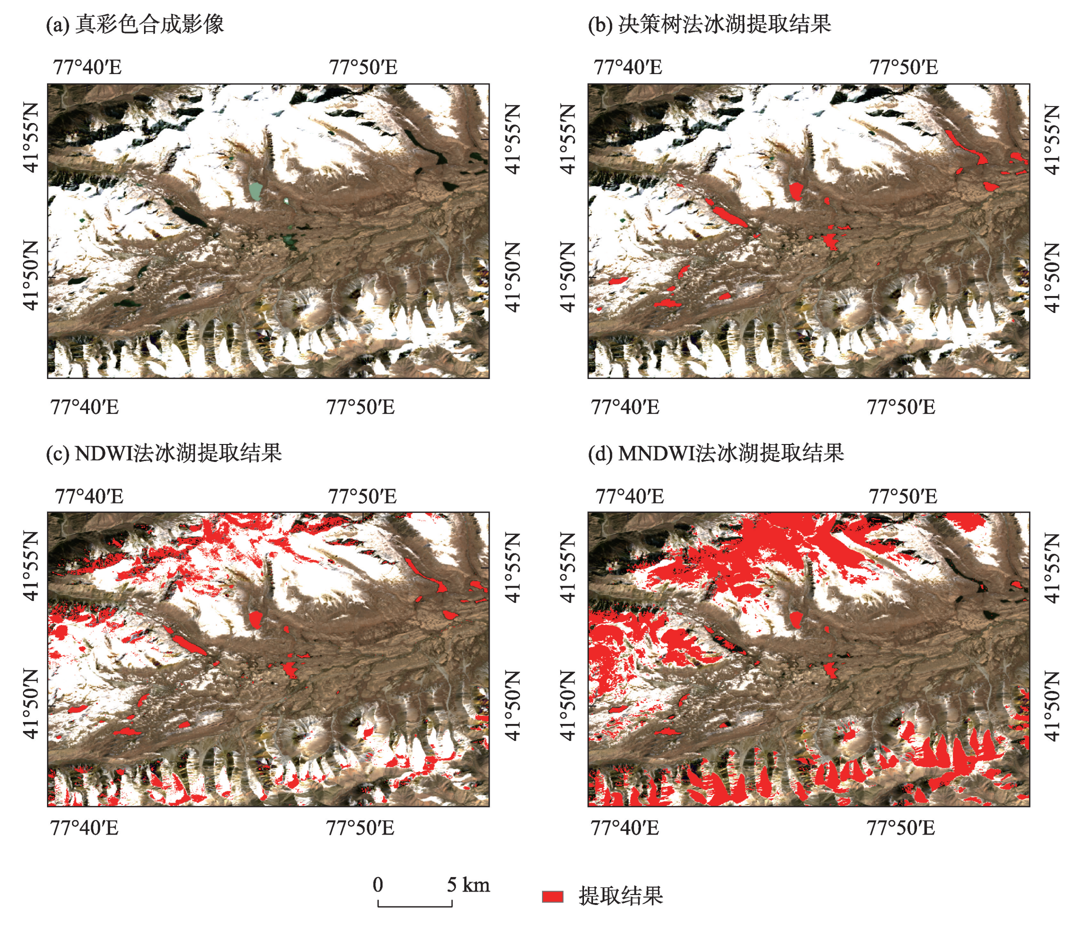


图4　冰湖提取结果对比

Fig. 4　Comparison of glacial lake extraction results

机选取了1750个冰湖样点和1750个非冰湖样点，导入冰湖提取结果中。经过混淆矩阵计算出生产者精度(PA)、用户精度(UA)、总体精度(OA)和Kappa系数作为天山地区冰湖精度评价的依据，结果如表1所示。在混淆矩阵计算后冰湖提取的总体精度为89.14%,Kappa系数为0.783,生产者精度及用户精度也处于较高水平。考虑到Kappa系数的意义难以解释的问题[40]，因此，加入F1分数进行精度验证补充，冰湖的F1分数为87.85%(表2)。

表1　天山冰湖提取精度混淆矩阵

Tab. 1　Tianshan glacial lake extraction accuracy confusion matrix

	验证样点		总计	用户精度/%
	冰湖	非冰湖		
冰湖	1374	4	1378	99.71
非冰湖	376	1746	2122	82.28
总计	1750	1750		总体精度:89.14%
生产者精度/%	78.51	99.77		Kappa系数:0.783

表2　天山冰湖提取精度

Tab. 2　Tianshan glacial lake extraction accuracy

准确率/%	召回率/%	F1分数/%
99.71	78.51	87.85

3.3　提取误差分析

本研究中建立了以高程、坡度、缓冲区范围、NDWI、SWI共5项规则的决策树法，但由于天山地区地形复杂，冰湖间的地形特征也存在差异。如图5所示冰湖中，对于湖中裸地部分可以剔除，但存在少部分冰湖信息未被提取，在对5项规则进行查看后发现，未识别部分通过了高程、缓冲区范围、NDWI、SWI共4项规则的筛选，但由于处于坡度较陡区域，坡度范围在10°~24°，因此，未通过坡度规则的筛选。

坡度规则的建立可以消除大部分山体阴影的影响，但由于冰湖地形特征不一，难以覆盖所有情况。从总体来看，坡度规则的建立对于冰湖提取精

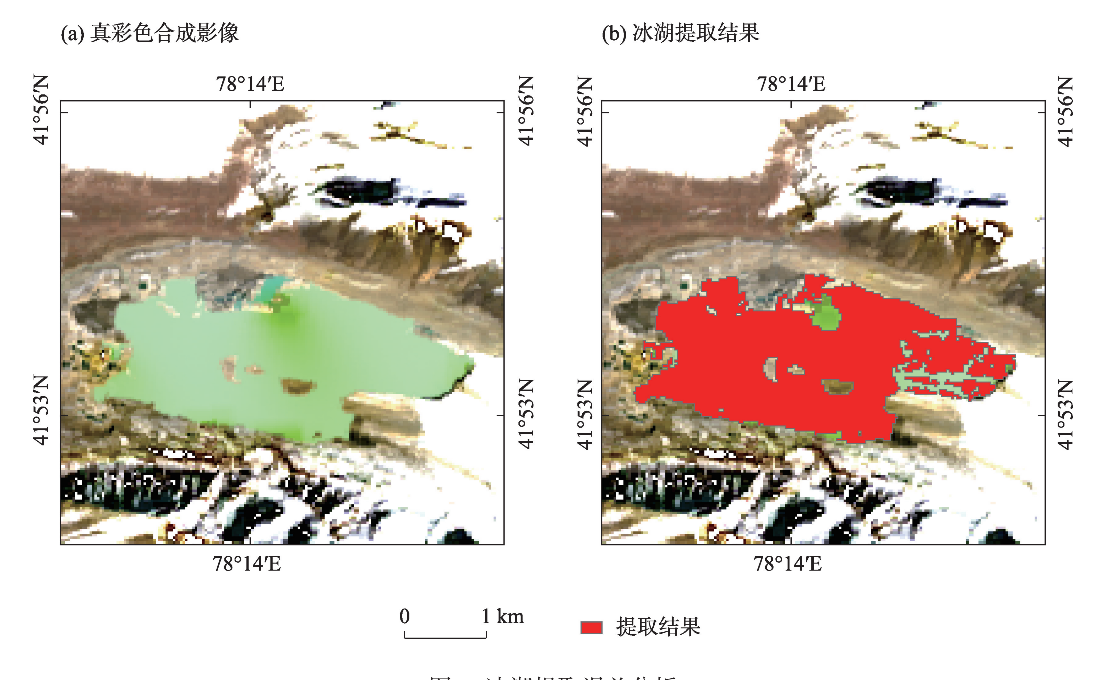


图5　冰湖提取误差分析

Fig. 5　Analysis of glacial lake extraction error

度的提升利大于弊。

4　结 论

本研究以Landsat 8地表反射率数据作为数据源，采用缨帽变换湿度指数提取冰川建立冰湖分布缓冲区，根据冰湖的空间位置（缓冲区范围）、地形特征（坡度、高程）以及光谱特征，构建了冰湖决策树提取方法，实现了对天山地区冰湖的精确提取，为冰湖监测、干旱区水资源调控提供一种方法选择。得出以下结论：

（1）通过对比分析决策树法、NDWI、MNDWI的提取结果后发现，NDWI法尽管能提取到大部分冰湖信息，但无法消除积雪和山体阴影的影响；MNDWI法的提取效果一般，不适用于复杂地形山区；本文构建的决策树法结合了冰湖空间位置、地形特征和光谱特征，能准确提取冰湖，并有效消除山体阴影和积雪的影响。

（2）对于复杂山区，单一的水体指数无法区分水体信息和山体阴影。决策树方法通过先验知识添加一系列规则，使用逻辑运算实现水体信息的有效提取和山体阴影、积雪影响的消除，有效弥补了传统水体指数受山体阴影影响较大的不足。

（3）基于Landsat卫星影像，决策树法可用于天山冰湖的快速提取，提取结果总体精度为89.14%，Kappa系数为0.783，F1分数为87.85%。结果具有较高可信度，为冰湖的动态监测与研究分析提供了一种较为高效的提取方法。

参考文献（References）：

[1] 黄启厅. 干旱区山地冰川—冰湖协同演变遥感监测与气候变化响应机制研究[D]. 北京: 中国科学院大学 (中国科学院遥感与数字地球研究所), 2017. [Huang Qiting. Study on Remotely Sensed Monitoring of the Synergic Evolution of Alpine Glacier-Glacier Lakes and Their Response Mechanism to the Climate Change in the Arid Areas[D]. Beijing: Institute of Remote Sensing and Digital Earth Chinese Academy of Sciences, 2017.]

[2] 李想. 基于GEE的近三十年亚洲高山区冰川雪线变化研究[D]. 西安: 西北大学, 2021. [Li Xiang. Research on Glacier Snow-line Changes in High Mountain Asia in the Past 30 Years Using GEE[D]. Xi'an: Northwest University, 2021.]

[3] 姚檀栋, 刘时银, 蒲健辰, 等. 高亚洲冰川的近期退缩及其对西北水资源的影响[J]. 中国科学(D辑), 2004, 34(6): 535-543. [Yao Tandong, Liu Shiyin, Pu Jianchen, et al. Recent retreat of high Asian glaciers and their impact on northwest water resources [J]. Science China Earth Sciences, 2004, 34(6): 535-543.]

[4] 邓海军, 陈亚宁. 中亚天山山区冰雪变化及其对区域水资源的影响[J]. 地理学报, 2018, 73(7): 1309-1323. [Deng Haijun, Chen Yaning. The glacier and snow variations and their impact on water resources in mountain regions: A case study in Tianshan Mountains of Central Asia[J]. Acta Geographica Sinica, 2018, 73(7): 1309-1323.]

[5] 王宗太. 新疆的冰川[J]. 干旱区地理, 1991, 14(1): 18-24. [Wang Zongtai. The glaicer of Xinjiang[J]. Arid Land Geography, 1991, 14 (1): 18-24.]

[6] 姚晓军, 刘时银, 韩磊, 等. 冰湖的界定与分类体系——面向冰湖编目和冰湖灾害研究[J]. 地理学报, 2017, 72(7): 1173-1183. [Yao Xiaojun, Liu Shiyin, Han Lei, et al. Definition and classification systems of glacial lake for inventory and hazards study[J]. Acta Geographica Sinica, 2017, 72(7): 1173-1183.]

[7] 陈晨, 郑江华, 刘永强, 等. 近20年中国阿尔泰山区冰川湖泊对区域气候变化响应的时空特征[J]. 地理研究, 2015, 34(2): 270-284. [Chen Chen, Zheng Jianghua, Liu Yongqiang, et al. The response of glacial lakes in the Altay Mountains of China to climate change during 1992-2013[J]. Geographical Research, 2015, 34 (2): 270-284.]

[8] 毕海芸, 王思远, 曾江源, 等. 基于TM影像的几种常用水体提取方法的比较和分析[J]. 遥感信息, 2012, 27(5): 77-82. [Bi Haiyun, Wang Siyuan, Zeng Jiangyuan, et al. Comparison and analysis of several common water extraction methods based on TM lmage[J]. Remote Sensing Information, 2012, 27(5): 77-82.]

[9] Wang X, Guo X, Yang C, et al. Glacial lake inventory of high-mountain Asia in 1990 and 2018 derived from Landsat images[J]. Earth System Science Data, 2020, 12(3): 2169-2182.

[10] 骆剑承, 盛永伟, 沈占锋, 等. 分步迭代的多光谱遥感水体信息高精度自动提取[J]. 遥感学报, 2009, 13(4): 610-615. [Luo Jiancheng, Sheng Yongwei, Shen Zhanfeng, et al. Step by step iterative multi-spectral remote sensing water information extraction with high precision[J]. National Remote Sensing Bulletin, 2009, 13 (4): 610-615.]

[11] Govindha Raj B K, Kumar V K, SN R. Remote sensing-based inventory of glacial lakes in Sikkim Himalaya: Semi-automated approach using satellite data[J]. Geomatics, Natural Hazards and Risk, 2013, 4(3): 241-253.

[12] Mitkari K V, Arora M K, Tiwari R K. Extraction of glacial lakes in Gangotri glacier using object-based image analysis[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2017, 10(12): 5275-5283.

[13] 闫斌, 贾洪果, 任文静, 等. 基于NDWI-NDSI组合阈值法的布加岗日冰湖提取及其变化分析[J]. 遥感学报, 2022, 26(11): 2344-2353. [Yan Bin, Jia Hongguo, Ren Wenjing, et al. Glacier lake extraction and variation analysis of the Bujiagangri glacier based on the NDWI-NDSI combination threshold method[J]. National Remote Sensing Bulletin, 2022, 26(11): 2344-2353.]

[14] 陈方, 王金晓, 张美美, 等. 基于历史边界的喜马拉雅山脉冰湖提取方法对比研究[J]. 冰川冻土, 2023, 45(4): 1413-1427. [Chen Fang, Wang Jinxiao, Zhang Meimei, et al. Comparative study on the extraction methods of Himalayan glacial lakes based on historical boundaries[J]. Journal of Glaciology and Geocryology, 2023, 45(4): 1413-1427.]

[15] 李宇宸, 张军, 刘陈立. Sentinel-2影像的云南千湖山细小冰湖提取方法[J]. 测绘科学, 2021, 46(4): 114-120. [Li Yuchen, Zhang Jun, Liu Chenli. Extraction method of alpine small lacial lake in Qianhu Mountain area of Yunnan province based on Sentinel-2 image[J]. Science of Surveying and Mapping, 2021, 46(4): 114-120.]

[16] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//Medical image computing and computer-assisted intervention-MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18. Springer International Publishing, 2015: 234-241.

[17] Li J, Sheng Y. An automated scheme for glacial lake dynamics mapping using Landsat imagery and digital elevation models: A case study in the Himalayas[J]. International Journal of Remote Sensing, 2012, 33(16): 5194-5213.

[18] 潘梦, 曹云刚. 高亚洲地区冰湖遥感研究进展与展望[J]. 自然资源遥感, 2021, 33(1): 1-8. [Pan Meng, Cao Yungang. Present status and perspectives of remote sensing survey of glacial lakes in High Asia[J]. Remote Sensing for Natural Resources, 2021, 33(1): 1-8.]

[19] Jain S K, Sinha R K, Chaudhary A, et al. Expansion of a glacial lake, Tsho Chubda, Chamkhar Chu Basin, Hindukush Himalaya, Bhutan[J]. Natural Hazards, 2015, 75: 1451-1464.

[20] Veh G, Korup O, Roessner S, et al. Detecting Himalayan glacial lake outburst floods from Landsat time series[J]. Remote Sensing of Environment, 2018, 207: 84-97.

[21] Bhardwaj A, Singh M K, Joshi P K, et al. A lake detection algorithm (LDA) using Landsat 8 data: A comparative approach in glacial environment[J]. International Journal of Applied Earth Observation and Geoinformation, 2015, 38: 150-163.

[22] 吴坤鹏. 天山冰湖变化及其影响研究[D]. 湘潭: 湖南科技大学, 2014. [Wu Kunpeng. The Change of Glacial Lake and Its Influence in Tianshan Mountains[D]. Xiangtan: Hunan University of Science and Technology, 2014.]

[23] 车涛, 李新, Mool P K, 等. 希夏邦马峰东坡冰川与冰川湖泊变化遥感监测[J]. 冰川冻土, 2005, 27(6): 801-805. [Che Tao, Li Xin, Mool P K, et al. Monitoring glaciers and associated glacial lakes on the east slopes of mount Xixabangma from remote sensing images[J]. Journal of Glaciology and Geocryology, 2005, 27(6): 801-805.]

[24] 陈亚宁, 李稚, 方功焕, 等. 气候变化对中亚天山山区水资源影响研究[J]. 地理学报, 2017, 72(1): 18-26. [Chen Yaning, Li Zhi, Fang Gonghuan, et al. lmpact of climate change on water resources in the Tianshan Mountians, Central Asia[J]. Acta Geographica Sinica, 2017, 72(1): 18-26.]

[25] Gardelle J, Arnaud Y, Berthier E. Contrasted evolution of glacial lakes along the Hindu Kush Himalaya mountain range between 1990 and 2009[J]. Global and Planetary Change, 2011, 75(1-2): 47-55.

[26] Jain R M, Mody K, Mishra A, et al. Physicochemical characterization of biosurfactant and its potential to remove oil from soil and cotton cloth[J]. Carbohydrate Polymers, 2012, 89(4): 1110-1116.

[27] Richardson S D, Reynolds J M. An overview of glacial hazards in the Himalayas[J]. Quaternary International, 2000, 65: 31-47.

[28] 殷永胜, 王欣, 刘时银, 等. 1990—2020年中国冰湖变化特征及影响因素[J]. 湖泊科学, 2023, 35(1): 358-367. [Yin Yongsheng, Wang Xin, Liu Shiyin, et al. Characteristics and influence factors of the glacial lake changes in China from 1990 to 2020[J]. Journal of Lake Sciences, 2023, 35(1): 358-367.]

[29] Crist E P, Kauth R J. The tasseled cap de-mystified[J]. Photogrammetric Engineering and Remote Sensing, 1986, 52: 81-86.

[30] Crist E P. A TM tasseled cap equivalent transformation for reflectance factor data[J]. Remote sensing of Environment, 1985, 17(3): 301-306.

[31] 周远刚, 赵锐锋, 张丽华, 等. 博格达峰地区冰川和积雪变化遥感监测及影响因素分析[J]. 干旱区地理, 2019, 42(6): 1395-1403. [Zhou Yuangang, Zhao Ruifeng, Zhang Lihua, et al. Remote sensing monitoring of the change of glacier and snow cover and its influencing factors in Mount Bogda[J]. Arid Land Geography, 2019, 42(6): 1395-1403.]

[32] 王欣, 刘时银, 姚晓军, 等. 我国喜马拉雅山区冰湖遥感调查与编目[J]. 地理学报, 2010, 65(1): 29-36. [Wang Xin, Liu Shiyin, Yao Xiaojun, et al. Glacier lake investigation and inventory in the Chinese Himalayas based on the remote sensing data[J]. Acta Geographica Sinica, 2010, 65(1): 29-36.]

[33] 杨成德, 王欣, 魏俊锋, 等. 2015年中国西部冰湖编目数据集[J]. 中国科学数据(中英文网络版), 2018, 3(4): 36-44. [Yang Chengde, Wang Xin, Wei Junfeng, et al. A dataset of glacial lake inventory of West China in 2015[J]. China Scientific Data, 2018, 3(4): 36-44.]

[34] McFeeters S K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features[J]. International Journal of Remote Sensing, 1996, 17(7): 1425-1432.

[35] 陈文倩, 丁建丽, 李艳华, 等. 基于国产GF-1遥感影像的水体提取方法[J]. 资源科学, 2015, 37(6): 1166-1172. [Chen Wenqian, Ding Jianli, Li Yanhua, et al. Extraction of water information based on China-made GF-1 remote sense image[J]. Resources Science, 2015, 37(6): 1166-1172.]

[36] 丁占峰, 李大军. 基于ONDWI水体指数的鄱阳湖水域信息提取[J]. 安徽农业科学, 2015, 43(6): 348-350. [Ding Zhanfeng, Li Dajun. The water region extraction of the Poyang Lake based on the ONDWI[J]. Journal of Anhui Agricultural Sciences, 2015, 43(6): 348-350.]

[37] 都金康, 黄永胜, 冯学智, 等. SPOT卫星影像的水体提取方法及分类研究[J]. 遥感学报, 2001, 5(3): 214-219. [Du Jinkang, Huang Yongsheng, Feng Xuezhi, et al. Study on water bodies extraction and classification from SPOT image[J]. National Remote Sensing Bulletin, 2001, 5(3): 214-219.]

[38] 文广超, 李兴, 吴冰洁, 等. 基于Landsat影像的柴达木盆地湖泊提取方法[J]. 干旱区研究, 2022, 39(3): 774-786. [Wen Guangchao, Li Xing, Wu Bingjie, et al. An automatic method for delineating lake surfaces in Qaidam Basin using Landsat images[J]. Arid Zone Research, 2022, 39(3): 774-786.]

[39] 徐涵秋. 利用改进的归一化差异水体指数(MNDWI)提取水体信息的研究[J]. 遥感学报, 2005, 9(5): 589-595. [Xu Hanqiu. A study on information extraction of water body with the Modified Normalized Difference Water lndex (MNDWI)[J]. National Remote Sensing Bulletin, 2005, 9(5): 589-595.]

[40] Foody G M. Explaining the unsuitability of the kappa coefficient in the assessment and comparison of the accuracy of thematic maps obtained by image classification[J]. Remote Sensing of Environment, 2020, 239: 111630.

Research on the extraction method of Tianshan glacier lake based on decision tree

LI Mengfan[1], ZHENG Jianghua[1,2], QIAN Anliang[1], LI Jiahui[1], Adiljan PARHAT[1],
WANG Zhe[1,2], MA Lisha[1,2], WANG Nan[1,2]

(1. College of Geography and Remote Sensing Sciences, Xinjiang University, Urumqi 830017, Xinjiang, China;
2. Xinjiang Key Laboratory of Oasis Ecology, Xinjiang University, Urumqi 830017, Xinjiang, China)

Abstract: The Tianshan region, located in the middle of the Eurasian continent, is a major distribution area of modern glaciers, and its glacial meltwater has formed a large number of and widespread glacial lakes. Glacial lakes are vital indicators of climate change and an important source of surface and groundwater supply in arid and semi-arid regions of Northwest China. The impact of topographic factors and spectral characteristics of ground objects makes it difficult to distinguish between glacial lakes, mountain shadow, and snow cover when extracting glacial lakes from remote sensing images using a single water index. In this study, based on the Google Earth Engine platform and Landsat 8 remote sensing images as the data source, a decision tree extraction method of the glacial lake was constructed according to the topographic characteristics (slope, elevation, and buffer analysis) and spectral characteristics of the glacial lake. This method was compared with the NDWI and MNDWI threshold methods. Experimental results demonstrated that the decision tree method can effectively reduce the impact of mountain shadow and snow cover and accurately extract glacial lake information. The overall accuracy of the extraction results was 89.14%, the Kappa coefficient was 0.783, and the F1 score was 87.85%. The decision tree method, which combines spatial topographic features and spectral features, is a relatively efficient extraction method for dynamic monitoring and research analysis of glacial lakes.

Keywords: glacial lake extraction; decision tree; Google Earth Engine; Tianshan